Molecular Ecology (2007) **16**, 2701–2711 doi: 10.1111/j.1365-294X.2007.03349.x

Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (*Kryptolebias marmoratus*)

ANDREY TATARENKOV,* HONG GAO,† MARK MACKIEWICZ,‡ D. SCOTT TAYLOR,§ BRUCE J. TURNER¶ and JOHN C. AVISE*

**Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA, †Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA, ‡Department of Genetics, University of Georgia, Athens, GA 30602, USA, §Brevard County Environmentally Endangered Lands Program, Melbourne, FL 32940, USA, ¶Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA*

Abstract

We employ a battery of 33 polymorphic microsatellite loci to describe geographical population structure of the mangrove killifish (*Kryptolebias marmoratus*), the only vertebrate species known to have a mixed-mating system of selfing and outcrossing. Significant population genetic structure was detected at spatial scales ranging from tens to hundreds of kilometres in Florida, Belize, and the Bahamas. The wealth of genotypic information, coupled with the highly inbred nature of most killifish lineages due to predominant selfing, also permitted treatments of individual fish as units of analysis. Genetic clustering algorithms, neighbour-joining trees, factorial correspondence, and related methods all earmarked particular killifish specimens as products of recent outcross events that could often be provisionally linked to specific migration events. Although mutation is the ultimate source of genetic diversity in *K. marmoratus*, our data indicate that interlocality dispersal and outcross-mediated genetic recombination (and probably genetic drift also) play key proximate roles in the local 'clonal' dynamics of this species.

Keywords: dispersal, gene flow, heterozygosity, inbreeding, microsatellites, *Rivulus*

Received 12 October 2006; revision accepted 20 March 2007

Introduction

Hermaphroditism occurs in about 6% of animal species distributed across more than 20 phyla (Jarne 1995; Jarne & Auld 2006). However, fewer than 20% of the 142 hermaphroditic species studied to date have self-fertilization rates greater than 0.8, and most of these are pulmonate mollusks. The rarity of predominant selfing in animals other than mollusks makes it especially interesting to examine exceptional cases. Here we address the population genetic features of self-fertilization in the mangrove killifish, *Kryptolebias* (formerly *Rivulus*) *marmoratus*, the only vertebrate species known to self-fertilize routinely and also the only known vertebrate that shows androdioecy (the presence of males as well as hermaphrodites).

Correspondence: Andrey Tatarenkov, Fax: (949) 824 2181; E-mail: tatarenk@uci.edu

In *K. marmoratus*, each hermaphrodite normally fertilizes itself, with syngamy occurring inside an internal 'ovotestis' (Harrington 1963; Sakakura *et al*. 2006). Most populations contain hermaphrodites almost exclusively, but males are common (10–25% frequency) at some sites in Belize (Turner *et al*. 2006). Males can be generated in the laboratory by exposing developing embryos or immature hermaphrodites to temperature shifts, and they are also known to arise 'spontaneously' by loss of female reproduction function in adult hermaphrodites (Turner *et al*. 2006). However, factors that mediate male frequencies in nature are not understood.

As originally deduced from allograft experiments (Kallman & Harrington 1964; Harrington & Kallman 1968), populations of *K. marmoratus* typically consist of highly homozygous 'clones' that arise via the intense inbreeding that attends multigeneration selfing. Although intrapopulation variation was detected in early multilocus DNA fingerprinting studies, such genetic diversity was attributed

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

to mutation and migration (Turner *et al*. 1992), and outcrossing was thought to occur only at a Belize site that is rich in males (Lubinski *et al*. 1995; Taylor *et al*. 2001). However, recent microsatellite analyses have documented that outcrossing occurs routinely in *K. marmoratus* (Mackiewicz *et al*. 2006b, c), and that most populations have a mixed-mating system of predominant selfing with occasional outcrossing. At least some outcross events involve males (Mackiewicz *et al*. 2006a), which presumably shed sperm on unfertilized ova that are extruded occasionally by hermaphrodites. Courtship displays and spawning behaviours typical of other killifishes have been observed in *K. marmoratus*, with the hermaphrodite assuming the 'female' role (Kristensen 1970).

The euryhaline mangrove killifish has a broad distribution, extending from coastal mangrove forests of peninsular Florida through the Bahamas, most Caribbean islands, the Yucatan Peninsula, and the Atlantic coasts of Central and South America to southeastern Brazil (Davis *et al*. 1990). 'Reproductive assurance' might be one important factor underlying this species' apparent capacity to disperse and colonize, because even a single selfing colonist could in principle establish a new population. Another factor underlying colonizing propensity is a suite of amphibious life-history features of this species (and other aplocheilid killifishes) known as emersion. Adults can leave the water for up to 10 weeks (Taylor 1990), respiring and eliminating ammonia cutaneously in damp environments (Grizzle & Thiyagarajah 1987; Frick & Wright 2002; Litwiller *et al*. 2006), and they can travel overland by flipping and jumping. Such behaviour permits escape from locally elevated concentrations of hydrogen sulphide, a common toxic component of mangrove habitats (Abel *et al*. 1987). Emersion may also be involved in reproduction, because the few embryos discovered to date in the field were oviposited out of water (Taylor 1990, 2000). Open-ocean dispersal has not been observed directly, but plausibly occurs via floating mangrove litter. Adults typically inhabit the burrows of land crabs, cavities in mangrove material, or moist termite galleries in rotting logs (Davis *et al*. 1990). When moved by tides and wind, such fish-containing debris could disperse individuals over long distances. The fertilized ova are also well suited for waif dispersal because they can readily attach to leaves, twigs, or other debris that sometimes float to the sea. Embryos are protected against desiccation by a tough chorion membrane, and can survive out-of-water for prolonged periods, hatching readily when hydrated (Ritchie & Davis 1986).

Previous studies of geographical population structure in *K. marmoratus* have been rather uninformative due to a paucity of genetic variation at the molecular markers utilized to date. For example, Vrijenhoek (1985) found no variation at 31 allozyme loci within or among strains from Florida and the Netherlands Antilles; and with regard to mitochondrial DNA, Weibel *et al*. (1999) observed only two cyt *b* sequences in 13 specimens from Brazil, Belize, Bahamas, and Florida, as did Sato *et al*. (2002) for control region sequences in 10 specimens from these locales. Here we utilize a large battery of polymorphic microsatellite loci to assess spatial population structure and deduce possible connections between dispersal and outcrossing. Our data indicate that immigration events are followed at some point by outcrossing with natives, which can quickly generate vast new genotypic variety.

Materials and methods

Samples

Locations and sample sizes are detailed in Table 1 and Fig. 1. Specimens were caught by miniature traps inserted into crab holes, by dip-netting (details in Taylor *et al*. 2004), by miniature hook and line, or by seining (at the Charlotte County site). Samples from Twin Cays, Belize, collected in 1991 and 2005, were also used to test for possible temporal shifts (across more than a decade) in allele frequencies, genotypic composition, and outcrossing rates.

Microsatellites

We utilized 33 of the 36 microsatellite loci developed by Mackiewicz *et al*. (2006a). Loci R24 and R45 were not used here because they cover the same genomic regions as R9 and R7, respectively, and locus R36 was excluded because it was invariable across all surveyed populations.

Statistical analyses

To evaluate genetic relationships *among individuals*, we used distance estimates (*Dps* values) based on the proportions of shared alleles (Bowcock *et al*. 1994). Matrices of *Dps* for the original data set, and for 100 data sets produced by bootstrapping, were calculated in MSANALYSER (Dieringer & Schlötterer 2003). To find bootstrap support values, matrices for the bootstrapped samples were further processed through modules NEIGHBOUR and CONSENSUS of the PHYLIP package (Felsenstein 1993). Neighbour-joining (NJ) trees were constructed in MEGA (Kumar *et al*. 2004). To display the overall genotypic associations of individuals, a factorial correspondence analysis (FCA) was performed using the procedure implemented in GENETIX (version 4.04; Belkhir *et al*. 2003). Correspondent analysis identifies a series of orthogonal axes to identify trends that explain the data variation, with each subsequent axis explaining a decreasing amount of variation. The placement of individuals along these axes can be used to convey the degree of similarity of individuals (or the degree of difference between them).

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

Table 1 Summary of genetic variation within 14 *Kryptolebias marmoratus* collections based on 33 microsatellite loci

Population	Tag	Sample size	Proportion of polymorphic loci (95% criterion)	Mean no. of alleles per locus	Expected heterozygosity	Observed heterozygosity	F_{IS}
Charlotte County, Florida	CC	17	0.58	2.1	0.258	0.022	0.917
Marco Island, Florida	MI	8	0.82	2.7	0.381	0.068	0.831
Lostman's River, Florida	LM	6	0.55	2.3	0.354	0.005	0.987
Harney River	HR	10	0.73	2.8	0.404	0.003	0.993
Tarpon Bay	TB	12	0.70	2.7	0.396	0.003	0.994
Shark River, Florida	SR	8	0.64	2.6	0.361	0	1
Everglades Nat'l Park, Florida	EP	4	0.73	2.1	0.432	0	1
Long Key, Florida	LK	7	0.61	1.6	0.217	0.082	0.640
No Name Key, Florida	NK	5	0.85	2.4	0.466	0.170	0.663
St Lucie County, Florida	SL	12	0.85	3.6	0.495	0.081	0.842
Exuma Island, Bahamas	EI	12	0.33	1.4	0.135	0	1
San Salvador Is., Bahamas	SS	5	0.61	2	0.326	0.012	0.967
Twin Cays & Belize 1991	TC91	21	0.94	6.8	0.660	0.472	0.292
Twin Cays & Belize 2005	TC05	101	0.97	10.4	0.688	0.502	0.270

F_{IS} represents coefficient of inbreeding. All F_{IS} values are highly significant ($P < 0.0001$) as evaluated by randomization in FSTAT (Goudet 1995).

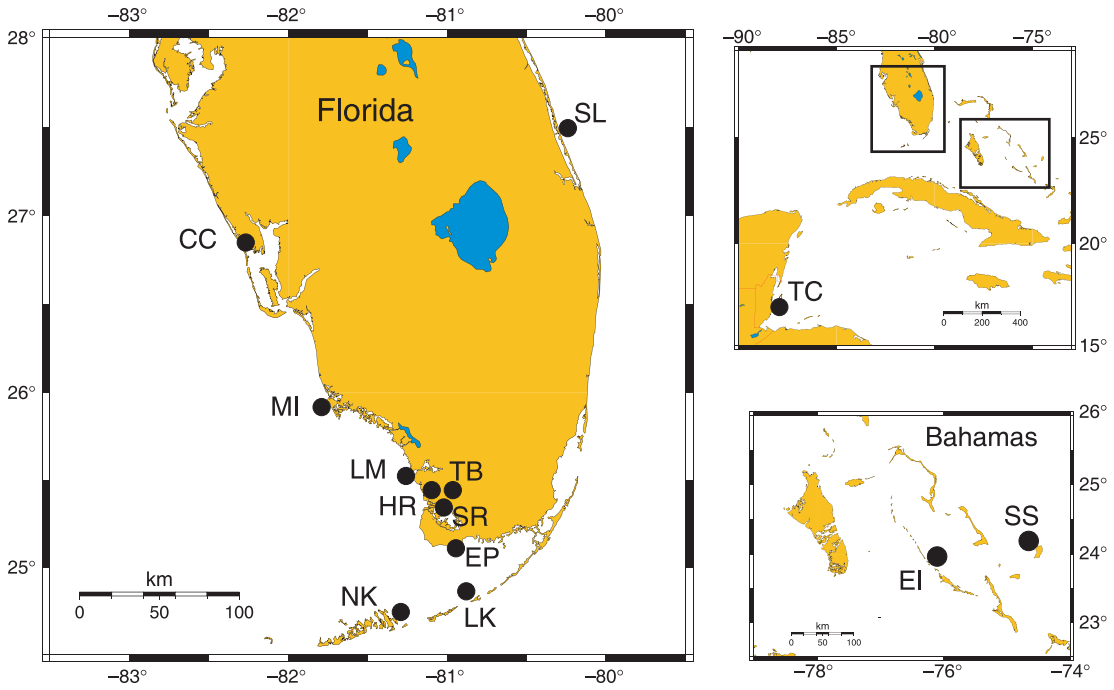

Fig. 1 Sample locations (populations are labelled as in Table 1).

To assess overall differentiation at the population level, we used FSTAT (version 2.9.3.2; Goudet 1995) to calculate F_{ST} and conduct exact *G*-tests based on 10 000 randomizations (Goudet *et al*. 1996). We typically randomized genotypes in these tests because all samples showed extensive departures from Hardy–Weinberg equilibrium. However, we also permuted alleles in some cases (such as when assessing possible temporal variation in allele frequencies at the Twin Cays site in Belize). To avoid confounding geographical with temporal variation, we excluded the 1991 Twin Cays sample from F_{ST} analyses. Also, because our sampling design involved no obvious hierarchy, we conducted the F_{ST} analyses separately for each set of populations of interest. The program FSTAT was also used to

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

measure departures from Hardy–Weinberg equilibrium. *P* values for F_{IS} for each locus were based on 46 200 randomizations, and *P* values over all loci were calculated from a weighted average of the statistic obtained for each locus.

Clustering of individuals with simultaneous estimates of selfing rates

We used the program STRUCTURE (version 2.1; Pritchard *et al*. 2000; Falush *et al*. 2003), which implements a model-based clustering algorithm that assigns individuals to a specified number (*K*) of clusters (presumed populations). Because the assignment is probabilistic, an individual may have joint membership in multiple populations, with membership coefficients summing to one. An important feature of this approach is that it again places emphasis on single individuals, thereby potentially allowing the identification of recent migration events as well as the overall degree of population admixture.

STRUCTURE in its original formulation was designed to infer structure among randomly mating subpopulations. The algorithm assumes that any observed deficit of heterozygotes in the whole sample is due to the Wahlund effect, and it attempts to minimize deviations from Hardy–Weinberg equilibrium (HWE) in subsamples for a given *K*. However, self-fertilization also leads to heterozygote deficiency, so to disentangle the two effects, it may be desirable to infer population structure and self-fertilization simultaneously. Such simultaneous estimation is achieved in the program INSTRUCT (Gao *et al*. 2007). INSTRUCT is an extension of STRUCTURE that incorporates estimates of selfing rates into the Bayesian clustering algorithm of STRUCTURE in an admixture context.

We ran INSTRUCT assuming 'independent allele frequencies' for five independent chains for each *K* value from 1 to 12, with each chain consisting of 1 million iterations with 500 000 burn-in iterations and 10 thinning intervals. The Gelman-Rubin statistic, implemented in INSTRUCT, was applied to assess convergence between chains. Runs with the highest *LnP(D)* probability are reported for each *K*. Graphical outputs from INSTRUCT were produced with help from the program DISTRUCT (Rosenberg *et al*. 2002).

To evaluate the importance of accounting for selfing rates on assessments of geographical structure, we also conducted the same analyses using STRUCTURE, and then compared results. We explored population structure using admixture models that assumes 'correlated allele frequencies' as well as 'independent allele frequencies'.

Results

Levels of variation

Thirty-three microsatellite loci were highly variable and thus potentially informative about population structure (Tables 1 and 2). Numbers of alleles ranged from two (at locus B86) to 55 (at locus R10), and averaged 13.7 per locus. Gene diversity in the total population sample ranged from 0.13 (at B86) to 0.95 (at R10), with an overall estimate of 0.64. Values for the fixation index (updated from Mackiewicz *et al*. 2006b; after exclusion of the three loci mentioned above) are compiled in Table 1. All populations showed statistically significant heterozygote deficits, as determined by permuting alleles separately for each locus and sample.

Table 2 Variability at 33 microsatellite loci in *Kryptolebias marmoratus* (total data set). Estimates of heterozygosities and F_{ST} are according to Nei (1987) (pp. 164–165) as implemented in FSTAT (Goudet 1995)

Locus	*N*	N_A	H_O	H_S	H_T	F_{ST}
R1	228	5	0.049	0.196	0.281	0.303
R3	228	22	0.134	0.495	0.804	0.385
R4	228	13	0.126	0.515	0.801	0.357
R5	225	11	0.087	0.644	0.849	0.241
R6	218	5	0.071	0.193	0.26	0.26
R7	228	6	0.086	0.194	0.245	0.208
R9	228	5	0.055	0.288	0.439	0.345
R10	207	55	0.101	0.67	0.948	0.293
R11	226	18	0.154	0.671	0.885	0.242
R16	228	13	0.115	0.483	0.756	0.361
R17	225	14	0.129	0.655	0.861	0.24
R18	227	15	0.133	0.579	0.774	0.252
R19	226	13	0.151	0.654	0.844	0.225
R22	226	21	0.104	0.663	0.892	0.257
R23	225	11	0.099	0.59	0.835	0.293
R25	228	15	0.153	0.451	0.743	0.394
R26	227	10	0.101	0.415	0.675	0.385
R27	227	13	0.112	0.589	0.888	0.337
R28	227	12	0.08	0.306	0.523	0.414
R30	227	10	0.127	0.515	0.795	0.353
R33	228	6	0.096	0.222	0.425	0.477
R34	228	6	0.046	0.11	0.318	0.653
R35	228	14	0.109	0.264	0.326	0.19
R37	226	37	0.152	0.825	0.946	0.128
R38	228	20	0.114	0.628	0.882	0.288
B10	151	12	0.078	0.367	0.736	0.502
B86	168	2	0	0	0.133	1
R86	222	15	0.103	0.204	0.459	0.555
R90	203	11	0.118	0.243	0.468	0.482
R92	211	5	0.095	0.193	0.23	0.163
R93	211	14	0.108	0.454	0.714	0.364
R103	221	7	0.035	0.241	0.643	0.625
R112	172	14	0.126	0.406	0.611	0.336
Overall	218.4	13.7	0.101	0.422	0.636	0.337

N, number of individuals; N_A, number of alleles; H_O, average observed heterozygosity per sample; H_S, average expected heterozygosity (gene diversity) per sample; H_T, gene diversity of the whole data set; F_{ST}, coefficient of genetic differentiation. All F_{ST} values are significant ($P < 0.01$ after Bonferroni correction) as determined by exact *G*-tests based on 10 000 randomizations of genotypes in FSTAT (Goudet 1995).

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

Expected heterozygosities in the Florida and Bahamas collections were always lower than those in Belize. For example, even when all 65 individuals from seven geographical sites (CC to EP in Table 1) along the south and west coast of Florida were pooled, heterozygosity (± SD) in this assemblage (0.494 ± 0.056) remained significantly lower than the corresponding estimate (0.688 ± 0.038) for the 2005 Twin Cays collection ($P < 0.01$, one-tailed t-test). This was true despite the fact that the latter estimate applies to a single population sampled at one time from one small island.

Cluster analysis

Classifications of individuals by the algorithms of INSTRUCT and STRUCTURE were very similar qualitatively. Below we describe results from INSTRUCT, but this description also closely fits results from STRUCTURE. At $K = 2$, two clusters of individuals were apparent, one consisting almost exclusively of Belize specimens (plus specimens NK1, NK3, and MI6 from Florida), and the other consisting exclusively of the remaining Florida and Bahamas specimens (Fig. 2). At $K = 3$, a new split within the Florida–Bahama assemblage distinguished two subgroups: most individuals from five locations in south and southwest Florida, plus NK in the Florida Keys, plus SL in eastern Florida and SS in the Bahamas, vs. most specimens from CC and MI in western Florida, plus LK in the Florida Keys and EI in the Bahamas. At $K = 4$, the latter cluster split further into western Florida on the one hand and EI and LK on the other (although the position of LK specimens was ambiguous because some runs grouped them instead with western Florida). At $K = 5$, an indication of population structure emerged within the 1991 Twin Cays sample: six specimens were assigned to a TC subgroup distinct from all other Belize individuals collected in 1991 and 2005. (Additionally, LK grouped with western Florida rather than with EI. Overall, the affinity of LK seemed to fluctuate between these two regions.) At $K = 6$, SS and SL separated from the more inclusive cluster consisting mostly of individuals from southwest Florida (plus some from NK in the Florida Keys). Level $K = 7$ produced no new clusters corresponding to geographical populations, but instead introduced some heterogeneity into NK and EP individuals.

At increasingly large K values (from 8 to 12), additional clusters (not shown) appeared, occasionally corresponding to specific geographical locales but more often indicating that various sites in southern Florida are genetically heterogeneous at these finer levels. In particular, specimens from EP in southern Florida represent a mixture of diverse genetic backgrounds (an outcome already evidenced at $K = 6$). This finer genetic heterogeneity is further evidenced by INSTRUCT analyses (at $K = 2$–4) as applied specifically to the LM, SR, TB, and SR locations in southwestern Florida (not shown). In general, however, the utility of INSTRUCT (and STRUCTURE) at levels of $K > 6$ diminished as the subdivisions became finer and the programs began simply to identify each inbred line as a separate population.

Having considered the results of specimen clustering for populations as a whole, it is also instructive to highlight several specific cases (see arrows on Fig. 2). As mentioned, two individuals from NK in Florida appeared to group genetically with the Twin Cays population, an assignment that recurred at all monitored levels of K. Similarly, individual MI6 from Florida displayed a high coefficient of genetic membership with six individuals in the TC91 collection. But, unlike the two NK specimens whose genetic makeup grouped entirely in the Twin Cays cluster, MI6 had an apparent mixture of local (MI) and distant (TC) genetic diversity.

In addition to these striking cases, several less conspicuous but nonetheless intriguing genetic connections were apparent. At higher K levels, individuals CC4 and LK3 from western Florida and the Keys, respectively, had 50% genetic membership in the southwest Florida group (the other half being of local origin). Specimens SL8 from eastern Florida and MI8 from MI in western Florida also had high genetic connection to the southwest Florida group.

Neighbour-joining and FCA analyses

Results in the NJ tree (depicting individuals as units of analysis) were generally congruent with those from INSTRUCT and STRUCTURE (Fig. 3). In most cases (with the same few notable exceptions), individuals from particular sites grouped together genetically. Individuals from Twin Cays, for example, formed a well-defined cluster that also included the two atypical NK individuals. Similarly, the six TC91 individuals distinguished by INSTRUCT also formed a well-supported NJ cluster that included MI6. This cluster was not separate from the remaining Belize (TC) individuals but instead was deeply embedded within that assemblage. In another example, specimens from the Bahamas again formed two groups corresponding to the particular islands (Exuma or San Salvador). Likewise, with the exception of one individual in each case, specimens from SL, LK, CC, LM, and MI formed clusters corresponding to their respective localities. Finally, individuals from the Shark and Harney Rivers, and Tarpon Bay (SR, HR, and TB) did not cluster strictly according to location, but instead formed a broad group that also incorporated three individuals from NK, one specimen from LK, two EP individuals, and one specimen from LR. Additionally, as before, CC4, SL8, and MI8 did not group with other specimens in their respective geographical populations.

The FCA analysis (Fig. 4) provides yet another perspective on the relationships among individuals, and generally supports results obtained by INSTRUCT, STRUCTURE and NJ.

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

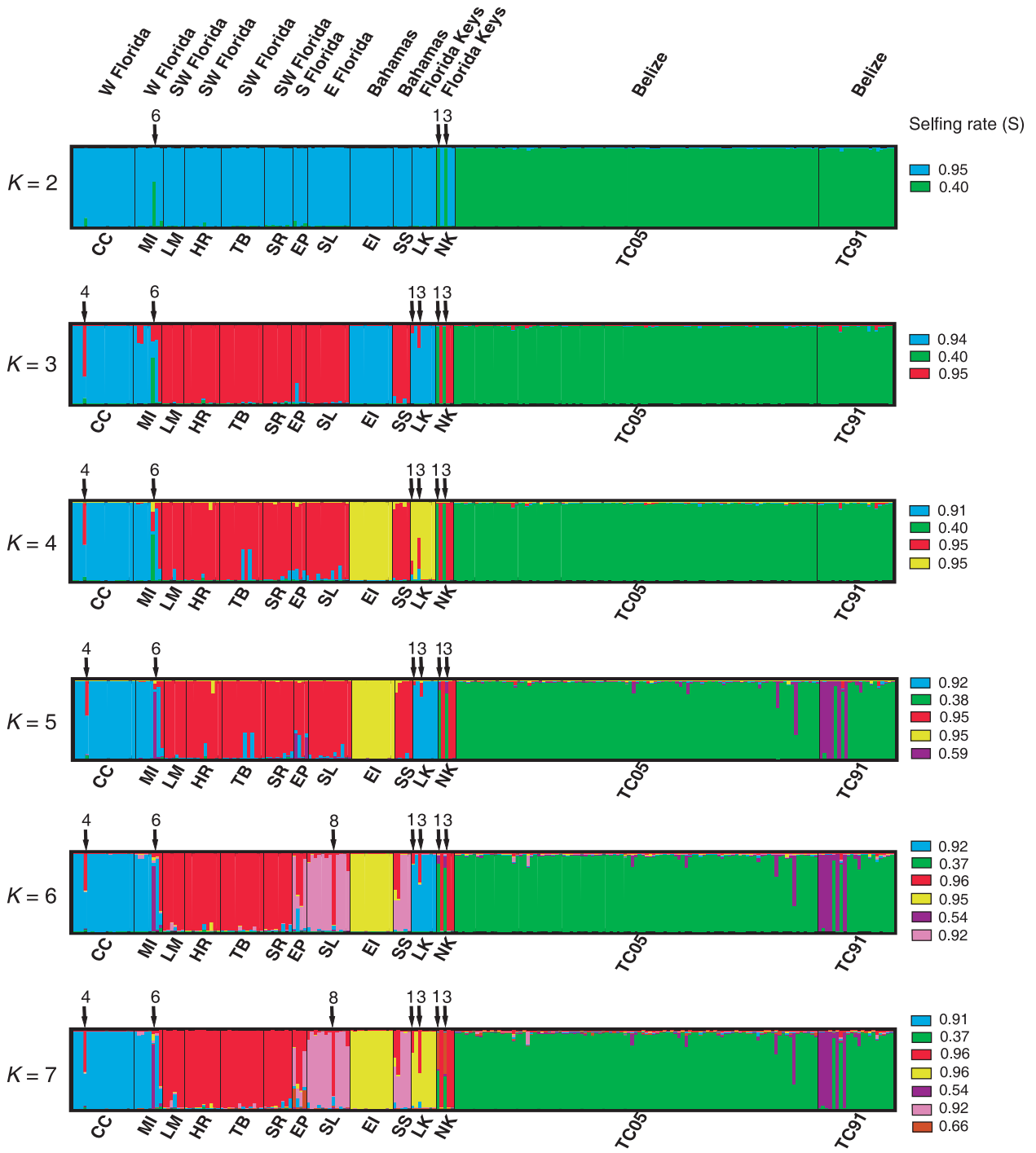

Fig. 2 Population structure assessed by genotypic clustering in INSTRUCT. Each killifish specimen is represented by a thin bar, often partitioned into coloured segments each representing an individual's proportionate genetic membership in a given *K*th cluster. Black lines separate sample locations. Specific sample sites (coded as in Table 1) are indicated below each panel and general geographical locations are shown at the top. The panels depict the highest probability outcomes (under the model assuming independent allele frequencies) at each level of *K* indicated to the left. Selfing rates (S), as determined by INSTRUCT, are shown to the right. Arrows indicate individuals discussed explicitly in the text.

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

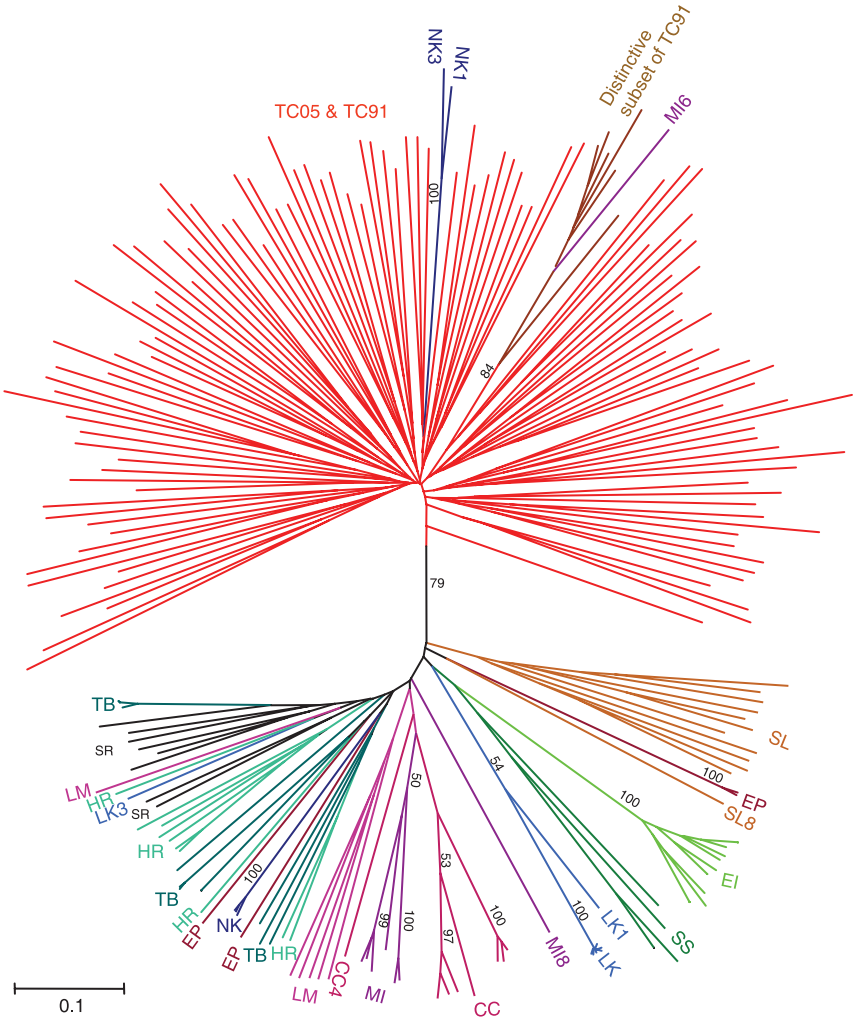

Fig. 3 Neighbour-joining tree for all 228 *Kryptolebias marmoratus* individuals sampled. The tree is constructed from *Dps* distances based on proportions of shared alleles. Bootstrap support is calculated from 100 replicates; values above 50% are shown on the nodes. Locations are labelled as in Table 1. Specimens discussed explicitly in the text are indicated by numbers following the population label.

One major split occurs along axis 1 between Twin Cays populations and the populations from Florida and Bahamas. Similarly to other analyses, individuals NK1 and NK3 are firmly embedded among TC specimens. Individual MI6 occupies an intermediate position between Twin Cays and the Floridian groupings, suggesting that its ancestry originates in both areas. Another split distinguishes Exuma Island individuals from those along the rest of axis 2. The remaining individuals form overlapping clouds, which is not surprising considering that the first two factorial axes explain a relatively low percentage of the total variation (or inertia). Apparently, as gauged by 450 alleles, the pattern of differentiation among samples is highly heterogeneous, such that the first two factorial axes are far from sufficient to account for the full complexity of the data.

Traditional analyses of geographical variation

The large number of loci screened allowed us to explore details of population structure at the level of single-specimen genetic affinities (above). Nonetheless, results of more traditional analyses (notably *F*-statistics; Hedrick 1999) that use composite population allele frequencies are also useful, not least because they permit convenient comparisons with other studies.

Genetic differentiation among geographical locales was pronounced and highly significant at many spatial scales (Table 3). For example, F_{ST} across the entire array of sites was 0.33 ($P < 0.001$), and it was 0.39 ($P < 0.001$) between Floridian and Bahamian samples. Even at the smallest spatial scale examined (about a dozen kilometres between HR, SR, and TB), genetic differentiation was substantial ($F_{ST} = 0.13$, $P < 0.001$).

Temporal variation in Twin Cays

Heterozygosities in the TC91 and TC05 samples were remarkably similar ($H_E \pm$ SD: 0.66 ± 0.04 and 0.69 ± 0.04, respectively; and $H_O \pm$ SD: 0.47 ± 0.02 and 0.50 ± 0.01, respectively). However, the two samples differed

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

Table 3 F_{ST} for various geographical levels in *Kryptolebias marmoratus*

Region	Samples included	F_{ST}
Caribbean, Florida, Bahamas	All but TC91	0.333*
Florida and Bahamas	All but TC91 and TC05	0.394*
Florida (including Keys)	CC,MI,LM,TB,SR,HR,EP,NK,LK,SL	0.311*
West Florida	CC,MI,LM,TB,SR,HR	0.282*
Basin of Shark and Harney Rivers	TB,SR,HR	0.132*

*$P < 0.001$.

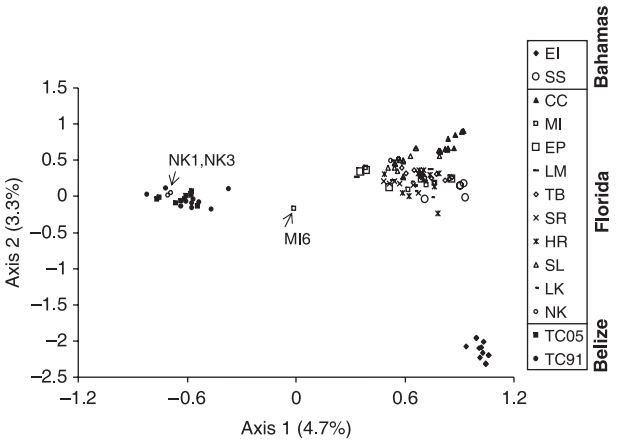

Fig. 4 Results of factorial correspondence analysis (FCA) showing the genetic relationships of *Kryptolebias marmoratus* individuals. Axes 1 and 2 are the first and second principle factors of variability that explain the indicated percentages of total genetic variation. All individuals were included in the FCA, but to avoid overcrowding in the figure only 10 representative specimens each from TC05 and TC91 are depicted here. Arrows indicate specific individuals discussed in the text.

Table 4 Temporal genetic variation and subdivision in *Kryptolebias marmoratus* from Twin Cays, Belize. TC05 is a sample ($N = 101$ specimens) collected in 2005; TC91 ($N = 21$) is from the same location sampled in 1991. Also analysed here are two distinct genetic subsets in the TC91 sample as identified in the INSTRUCT program. 95% confidence intervals (CI) for F_{ST} were obtained by bootstrapping loci (10 000 replicates) using the program GDA (Lewis & Zaykin 2001)

	F_{ST}	95% CI
TC05 vs. TC91 ($N = 21$)	0.023	[0.012; 0.034]
TC05 vs. TC91 ($N = 15$)	0.002	[–0.005; 0.010]
TC05 vs. TC91 ($N = 6$)	0.204	[0.156; 0.248]
TC91 ($N = 15$) vs. TC91 ($N = 6$)	0.214	[0.158; 0.265]

significantly in allelic frequencies at 19 of 33 polymorphic loci, and in genotype frequencies at 11 loci (in both cases, $P < 0.001$ in combined Fisher's test across loci). Nonetheless, the overall magnitude of genetic differentiation was relatively low ($F_{ST} = 0.023$; Table 4).

The cluster analysis in INSTRUCT had suggested hidden population subdivision in the TC91 sample (with six individuals appeared distinct from all others in the TC91 and TC05 Twin Cays collections). Accordingly, we subdivided TC91 in the F_{ST} analysis to ask whether these two subsets of specimens differ in allele and genotype frequencies. They do, at 19 loci (including R10, with a fixed allelic difference). The absolute level of divergence between these two subsets of specimens ($F_{ST} = 0.214$) also far exceeded the divergence ($F_{ST} = 0.002$) observed between the other TC91 specimens and those from TC05 (this latter difference was nonetheless statistically significant overall; Fisher's combined $P < 0.001$). The difference between TC05 and TC91 (even after excluding the six distinct individuals from TC91) indicates significant shifts in allelic frequency at this location between 1991 and 2005. We detected no change, however, in observed levels of heterozygosity between these two sampling periods.

The unusually dramatic genetic subdivision within the TC91 sample might be explained either by the presence of substantial genetic variation at this microgeographical scale or by a pre-1991 addition of migrants to that site. In any event, the maintenance of similar levels of observed heterozygosity in the two Belize samples separated by 14 years indicates that outcrossing at the TC site is an ongoing process, thus supporting a supposition by Lubinski *et al*. (1995) and refuting an alternative hypothesis that heterozygotes may have resulted from a rare outcross event in 1 year.

Selfing rate analysis

When clustering individuals, INSTRUCT concurrently estimates selfing rates. These are shown in Fig. 2 for each cluster (colour), and they are mostly similar to traditional estimates of selfing based on F_{IS} (Mackiewicz *et al*. 2006b). Thus, for the Florida populations (excluding the Keys), selfing rates as estimated by F_{IS} ranged from 0.91 to 1.00, and those estimated using INSTRUCT ranged from 0.91 to 0.96. Individuals from Belize displayed much lower rates of self-fertilization (range 0.42–0.44 as estimated by F_{IS}, and 0.37–0.54 according to INSTRUCT). However, differences between selfing estimates from F_{IS} and INSTRUCT were evident in cases of suspected population admixture. Thus, population NK from the Florida Keys showed relatively low

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

selfing rates (0.80), compared to other Florida populations as estimated by F_{IS} (*c*. 0.80), whereas INSTRUCT interprets the NK sample to be an admixture of local individuals with high rates of selfing (S = 0.96) and inferred immigrants from a population with a lower selfing rate (S = 0.37).

Discussion

Turner *et al*. (1992) found *no* matches in complex DNA fingerprints for mangrove killifish collected in different years at specific Florida sites, and they invoked genetic drift and frequent migration events to account for such high clonal turnover. Our recent genetic documentation of occasional outcrossing in Floridian (and other) locations (Mackiewicz *et al*. 2006b, c) indicates, however, that genetic recombination also plays a key immediate role in clonal diversity and dynamics. *De novo* mutations and (perhaps) interlocality gene flow are of course the ultimate sources of genetic variation in any population, but outcrossing in the otherwise selfing populations of *Kryptolebias marmoratus* is undoubtedly the potent proximal force capable of quickly generating vast genotypic variety at particular locales.

As shown here, the polymorphic microsatellite markers previously employed to document outcrossing in *K. marmoratus* also document significant population genetic structure in this species at all spatial scales examined (from tens to hundreds of kilometres). More interestingly, this large battery of markers has also allowed us to provisionally identify the geneses of particular specimens and strains, including specific outcross events and possible interlocality migrations.

Genetic dissections of specific outcross and migration events

Table 5 provides a summary of the six most compelling cases of outcrossing and/or migration suggested by our genetic data. These cases will be elaborated below, but first an important caveat: our inferences about the hypothesized geographical origins of particular migrant genotypes can only be educated guesses, given that we have not exhaustively sampled the sites examined nor have we genetically surveyed all other potential geographical sources of immigrants, which must be vast.

Nonetheless, various outcross events were genetically unambiguous. For example, specimen LK1 is a 19-locus heterozygote with distinctive alleles otherwise carried in homozygous condition by LK3 on the one hand and by the remaining LK specimens on the other, so it is an obvious immediate product of outcrossing. LK3 in turn was an inbred descendant of an apparent outcross between specimens with genotypes characteristic of LK and southwestern Florida (so a recent interlocality dispersal event between these two south Florida sites is also genetically implicated). Similarly, specimen CC4 (another multilocus heterozygote) appears to be the first-generation product of an outcross event perhaps involving a southwestern Florida immigrant into the CC site.

Several events of long-range dispersal were also implicated by the genetic data, namely represented by individuals MI6, NK1, NK3, and SL8 (Table 5). MI6 from Florida may register an immigration event (at face value from Belize) followed by outcrossing with native Florida strains. This fish in question was heterozygous at multiple loci for alleles otherwise mostly confined to these separate regions. NK1 and NK3 may represent a case of recent immigration. They are strikingly different from other NK specimens in their allelic composition and in terms of being heterozygous at 13 loci each (other NK specimens are completely homozygous). NK1 and NK3 firmly fall within the Twin Cays cluster by all analyses, thus suggesting their origin in that general region. However, being multilocus heterozygotes but without any genetic trace of admixture with local stock, NK1 and NK3 may be either recent arrivals or perhaps products of outcrossing between migrants (which might have been ecologically or behaviourally isolated from the local fish). Closer genotypic inspection further suggests that NK1 and NK3 are possible siblings, and that they are perhaps one or two selfing generations removed from the original outcross event (assuming that the level of

Table 5 Natural history of evident outcross events often associated with specifiable migration events (all as provisionally evidenced by various of the genetic analyses)

Specimens	Natural history
CC4	A multilocus heterozygote, apparently the result of an outcross of CC × southwest Florida
MI6	A multilocus heterozygote, apparently the result of an outcross TC × southwest Florida with subsequent selfing probably for 1–2 generations
SL8	A recent migrant (no outcross) from southwest Florida to SL
LK3	A descendant of an outcross LK × southwest Florida with subsequent selfing. This line later backcrossed with an LK lineage to produce LK1
LK1	A multilocus heterozygote, apparently a first-generation outcross of LK3 × LK
NK1, NK3	Multilocus heterozygotes, apparently migrants from Belize, subsequently selfed perhaps 1–2 times; possible siblings

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

heterozygosity in the source population is similar to that in Twin Cays). Finally, at least one putative instance of a recent immigration event did not yet involve subsequent outcrossing. Specimen SL8 from eastern Florida is homozygous at loci for alleles that otherwise characterize fish in southwestern Florida.

Additional (but far more general) evidence for occasional gene movement involves individuals from the LM, TB, HR, and EP sites in southwestern Florida. These form a well-mixed genetic cluster implying that population genetic exchange on a scale of 50–100 km may be rather common (and/or that these sites shared a relatively recent polymorphic ancestry).

Are migrant and outcross events linked?

Outcrossing is detected most readily when it involves inbred strains that differ at multiple loci, which as we have shown is much more likely to be true for specimens from different geographical sites. An inherent bias thus exists in favour of observing an apparent association between dispersal and outcrossing, even if per-capita outcrossing among natives is perhaps equal or more common. This bias is further exaggerated because genetically detectable outcrossing following an immigration event could take place in any generation following the immigrant's arrival. This raises the broader issue of whether particular immigration events in *K. marmoratus* took place in the recent or distant past.

In a random mating population, the presence of an individual with a distinct multilocus profile immediately suggests recent immigration (because Hardy–Weinberg equilibrium is restored in one generation, and gametic-phase disequilibrium between unlinked loci mostly disappears within a few generations). In a mostly selfing population, however, the full integration of migrant genotypes into the genetic background of the local gene pool is much slower. Indeed, under strict selfing, a migrant lineage could persist indefinitely in recognizable form (with subsequent genetic changes due only to *de novo* mutations).

Although we can point with our data to some highly plausible instances of recent immigration (as described above), we cannot by hard criteria eliminate a competing hypothesis that a given local population has long been highly subdivided internally. For example, in theory one portion of the population might reproduce mostly by selfing and the other more so by outcrossing, in which case the outcrossed part would remain relatively homogeneous genetically whereas inbred lineages in the selfed portion could build up substantial genetic differences over time. Another such possible generator of long-term local population structure might be otherwise cryptic physical or social barriers to microspatial dispersal. However, such possibilities seem unlikely for three reasons. First, *K. marmoratus* must be an effective colonizing species as gauged by its wide geographical distribution that includes isolated Caribbean islands. Second, the long-term persistence of large numbers of distinct lineages at any site would seem unlikely given the modest local population sizes often suspected for this species (Harrington 1971; but see Taylor *et al*. 2004). Third and perhaps most telling, our genetic data give no hint that any local population (with the possible but unlikely exception of NK, as described earlier) is composed of highly discrete pools of fish with long-term and consistently maintained mating differences (i.e. one subset being mostly random-mating and the other selfing).

In summary, although many questions remain about various details of 'clonal' diversity and dynamics in *K. marmoratus*, there can now be no doubt that migration events and outcross-mediated genetic recombination have played key evolutionary roles in this mixed-mating species.

Acknowledgements

Our work was supported by funds from the University of California at Irvine. We thank William P. Davis, William Dunson, John Grizzle, and Carole McIvor for specimen collection. We also thank Carlos D. Bustamante and Scott Williamson for helpful comments on a draft of the manuscript. Belize specimens were collected with the support of the Smithsonian Institution's Caribbean Coral Reef Ecology Program (CCRE Contribution no. 796).

References

Abel DC, Koenig CC, Davis WP (1987) Emersion in the mangrove forest fish *Rivulus marmoratus*, a response to hydrogen sulfide. *Environmental Biology of Fishes*, **18**, 67–72.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2003) GENETIX, *Logiciel Sous Windows Pour la Génétique Des Populations*. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier, France.

Bowcock AM, Ruiz-Linares A, Tomfohrde J *et al*. (1994) High-resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.

Davis WP, Taylor DS, Turner BJ (1990) Field observations of the ecology and habits of mangrove Rivulus (*Rivulus marmoratus*) in Belize and Florida (Teleostei: Cyprinodontiformes: Rivulidae). *Icthyological Explorations of Freshwaters*, **1**, 123–134.

Dieringer D, Schlötterer C (2003) MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Felsenstein J (1993) PHYLIP *(Phylogeny Inference Package), Version 3.5c*. Department of Genetics. University of Washington, Seattle.

Frick NT, Wright PA (2002) Nitrogen metabolism and excretion in the mangrove killifish *Rivulus marmoratus*. II. Significant ammonia volatilization in a teleost during air-exposure. *Journal of Experimental Biology*, **205**, 91–100.

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd

Gao H, Williamson S, Bustamante CD (2007) An MCMC approach for joint inference of population structure and inbreeding rates from multi-locus genotype data. *Genetics*, in press.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J, Raymond M, de Meeus T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Grizzle JM, Thiyagarajah A (1987) Skin histology of *Rivulus ocellatus marmoratus*, apparent adaptation for aerial respiration. *Copeia*, **1987**, 237–240.

Harrington RW Jr (1963) Twenty-four hour rhythms of internal self-fertilization and of oviposition by hermaphrodites of *Rivulus marmoratus*. *Physiological Zoology*, **36**, 325–321.

Harrington RW Jr (1971) How ecological and genetic factors interact to determine when self-fertilizing hermaphrodites of *Rivulus marmoratus* change into functional secondary males, with a reappraisal of modes of intersexuality among fishes. *Copeia*, **1971**, 389–432.

Harrington RW Jr, Kallman KD (1968) The homozygosity of clones of the self-fertilizing hermaphroditic fish *Rivulus marmoratus* (Cyprinodontidae, Atheriniformes). *American Naturalist*, **102**, 337–343.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Jarne P (1995) Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genetical Research*, **65**, 193–207.

Jarne P, Auld JR (2006) Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. *Evolution*, **60**, 1816–1824.

Kallman KD, Harrington RW Jr (1964) Evidence for the existence of homozygous clones in the self-fertilizing hermaphroditic teleost *Rivulus marmoratus* Poey. *Biological Bulletin*, **126**, 101–114.

Kristensen I (1970) Competition in three cyprinodont fish species in the Netherlands Antilles. *Studies of the Fauna of Curacao and Other Caribbean Islands*, **119**, 82–101.

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

Lewis PO, Zaykin D (2001) *GDA (Genetic Data Analysis): Computer Program for the Analysis of Allelic Data*. Available at: http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.

Litwiller SL, O'Donnell M, Wright PA (2006) Rapid increase in the partial pressure of NH_3 on the cutaneous surface of air-exposed mangrove killifish, *Rivulus marmoratus*. *Journal of Experimental Biology*, **209**, 1737–1745.

Lubinski BA, Davis WP, Taylor DS, Turner BJ (1995) Outcrossing in a natural population of a self-fertilizing hermaphroditic fish. *Journal of Heredity*, **86**, 469–473.

Mackiewicz M, Tatarenkov A, Perry A *et al*. (2006a) Microsatellite documentation of outcrossing between highly inbred laboratory strains of the self-fertilizing mangrove killifish (*Kryptolebias marmoratus*). *Journal of Heredity*, **97**, 508–513.

Mackiewicz M, Tatarenkov A, Taylor DS, Turner BJ, Avise JC (2006b) Extensive outcrossing and androdioecy in a vertebrate species that otherwise reproduces as a self-fertilizing hermaphrodite. *Proceedings of the National Academy of Sciences, USA*, **103**, 9924–9928.

Mackiewicz M, Tatarenkov A, Turner BJ, Avise JC (2006c) A mixed-mating strategy in a hermaphroditic vertebrate. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **273**, 2449–2452.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Ritchie SA, Davis WP (1986) Evidence for embryonic diapause in *Rivulus marmoratus*: laboratory and field observations. *Journal of the American Killifish Association*, **19**, 103–108.

Rosenberg NA, Pritchard JK, Weber JL *et al*. (2002) Genetic structure of human populations. *Science*, **298**, 2381–2385.

Sakakura Y, Soyano K, Noakes DLG, Hagiwara A (2006) Gonadal morphology in the self-fertilizing mangrove killifish, *Kryptolebias marmoratus*. *Ichthyological Research*, **53**, 427–430.

Sato A, Satta Y, Figueroa F *et al*. (2002) Persistence of Mhc heterozygosity in homozygous clonal killifish, *Rivulus marmoratus*: implications for the origin of hermaphroditism. *Genetics*, **162**, 1791–1803.

Taylor DS (1990) Adaptive specializations of the cyprinodont fish, *Rivulus marmoratus*. *Florida Scientist*, **53**, 239–248.

Taylor DS (2000) Biology and ecology of *Rivulus marmoratus*: new insights and a review. *Florida Scientist*, **63**, 242–255.

Taylor DS, Davis WP, Turner BJ (2004) Grovelling in the mangroves: 16 years in pursuit of the cyprinodont fish *Rivulus marmoratus* on the Belize Cays. *Atoll Research Bulletin* **No**. **525**, 1–14.

Taylor DS, Fisher MT, Turner BJ (2001) Homozygosity and heterozygosity in three populations of *Rivulus marmoratus*. *Environmental Biology of Fishes*, **61**, 455–459.

Turner BJ, Elder JF Jr, Laughlin TF, Davis WP, Taylor DS (1992) Extreme clonal diversity and divergence in populations of a selfing hermaphroditic fish. *Proceedings of the National Academy of Sciences, USA*, **89**, 10 643–10 647.

Turner BJ, Fisher MT, Taylor DS, Davis WP, Jarrett BL (2006) Evolution of 'maleness' and outcrossing in a population of the self-fertilizing killifish, *Kryptolebias marmoratus*. *Evolutionary Ecology Research*, **8**, 1475–1486.

Vrijenhoek RC (1985) Homozygosity and interstrain variation in the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. *Journal of Heredity*, **76**, 82–84.

Weibel AC, Dowling TE, Turner BJ (1999) Evidence that an outcrossing population is a derived lineage in a hermaphroditic fish (*Rivulus marmoratus*). *Evolution*, **53**, 1217–1225.

Andrey Tatarenkov, currently a postdoctoral research associate in the Avise laboratory, is interested in the use of molecular markers to study population genetics, incipient speciation, and phylogenetics. Hong Gao is a Ph.D. student employing statistical modeling in population genetics. Mark Mackiewicz did his Ph.D. thesis in Avise's lab on mating systems in fishes; he is currently a Postdoctoral Fellow at the National Cancer Institute in Bethesda. D. Scott Taylor is an ecologist interested in conservation, and Bruce J. Turner studies the evolutionary biology of fishes. John C. Avise has broad interests in molecular ecology, natural history, and evolution.

© 2007 US Federal Government
Journal compilation © 2007 Blackwell Publishing Ltd